



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
H04N 7/26, 7/36

A1

(11) International Publication Number: WO 99/43162

(43) International Publication Date: 26 August 1999 (26.08.99)

(21) International Application Number: PCT/US99/02318

(22) International Filing Date: 2 February 1999 (02.02.99)

(30) Priority Data:
09/027,140 20 February 1998 (20.02.98) US

(71) Applicant (for all designated States except US): THOMSON CONSUMER ELECTRONICS [US/US]; 10330 North Meridian Street, Indianapolis, IN 46290 (US).

(72) Inventor; and

(75) Inventor/Applicant (for US only): GOLIN, Stuart, Jay [US/US]; 626 Dutch Neck Road, East Windsor, NJ 08520 (US).

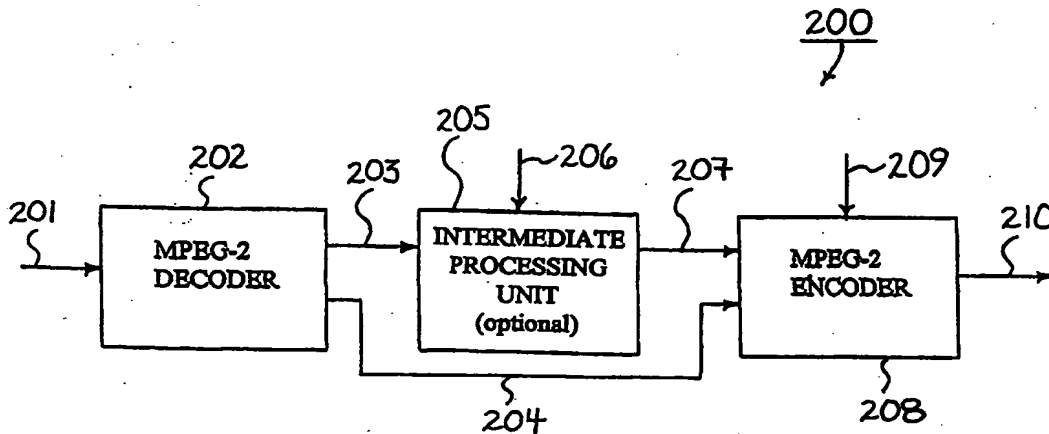
(74) Agent: CARTER, Jeffrey, D.; Thomson Consumer Electronics, P.O. Box 1976, INH 370, Indianapolis, IN 46206-1976 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

With international search report.

(54) Title: MOTION VECTOR EXTRAPOLATION FOR TRANSCODING VIDEO SEQUENCES



(57) Abstract

A transcoding method is performed by receiving a first bitstream of compressed image data having identifiable coding parameters. These parameters may relate to the GOP structure of pictures represented in the first bitstream, the size of the pictures represented in the first bitstream, whether the pictures represented in the first bitstream are field or frame pictures, and/or whether the pictures represented in the first bitstream define a progressive or interlaced sequence. First motion information is obtained from the first bitstream, and is used to extrapolate second motion information for a second bitstream of compressed image data. The second bitstream, which has one or more parameters different from the parameters of the first bitstream, is provided as a transcoded output.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| | | | | | | | |
|----|--------------------------|----|--|----|--|----|--------------------------|
| AL | Albania | ES | Spain | LS | Lesotho | SI | Slovenia |
| AM | Armenia | FI | Finland | LT | Lithuania | SK | Slovakia |
| AT | Austria | FR | France | LU | Luxembourg | SN | Senegal |
| AU | Australia | GA | Gabon | LV | Latvia | SZ | Swaziland |
| AZ | Azerbaijan | GB | United Kingdom | MC | Monaco | TD | Chad |
| BA | Bosnia and Herzegovina | GE | Georgia | MD | Republic of Moldova | TG | Togo |
| BB | Barbados | GH | Ghana | MG | Madagascar | TJ | Tajikistan |
| BE | Belgium | GN | Guinea | MK | The former Yugoslav Republic of Macedonia | TM | Turkmenistan |
| BF | Burkina Faso | GR | Greece | ML | Mali | TR | Turkey |
| BG | Bulgaria | HU | Hungary | MN | Mongolia | TT | Trinidad and Tobago |
| BJ | Benin | IE | Ireland | MR | Mauritania | UA | Ukraine |
| BR | Brazil | IL | Israel | MW | Malawi | UG | Uganda |
| BY | Belarus | IS | Iceland | MX | Mexico | US | United States of America |
| CA | Canada | IT | Italy | NE | Niger | UZ | Uzbekistan |
| CF | Central African Republic | JP | Japan | NL | Netherlands | VN | Viet Nam |
| CG | Congo | KE | Kenya | NO | Norway | YU | Yugoslavia |
| CH | Switzerland | KG | Kyrgyzstan | NZ | New Zealand | ZW | Zimbabwe |
| CI | Côte d'Ivoire | KP | Democratic People's Republic of Korea | PL | Poland | | |
| CM | Cameroon | KR | Republic of Korea | PT | Portugal | | |
| CN | China | KZ | Kazakhstan | RO | Romania | | |
| CU | Cuba | LC | Saint Lucia | RU | Russian Federation | | |
| CZ | Czech Republic | LI | Liechtenstein | SD | Sudan | | |
| DE | Germany | LK | Sri Lanka | SE | Sweden | | |
| DK | Denmark | LR | Liberia | SG | Singapore | | |
| EE | Estonia | | | | | | |

TITLE OF THE INVENTION**MOTION VECTOR EXTRAPOLATION FOR TRANSCODING
VIDEO SEQUENCES****GOVERNMENT LICENSE RIGHTS**

The United States Government has a paid-up license in this invention and the right in limited circumstances to require the patent owner to license others on reasonable terms as provided for by the terms of contract number 70NANBSH1171 awarded by the National Institute of Standards and Technology.

FIELD OF THE INVENTION

The present invention generally relates to motion estimation in the transcoding of a video sequence, and more particularly, to a technique for transcoding a received video sequence by using extrapolated motion information from the received video sequence, thereby minimizing or eliminating motion estimation from the transcoding process.

BACKGROUND OF THE INVENTION

Data compression techniques enable large quantities of data to be transmitted over relatively small bandwidths. The algorithms used in a compression system depend on the available bandwidth or storage capacity, the features required by the application, and the affordability of the hardware required for implementation of the compression algorithm (encoder as well as decoder). The moving pictures expert group-type 2 ("MPEG-2")

compression standard, which is hereby incorporated by reference, is a well-known technique for compressing video data. Motion estimation, which is the process video encoders use for calculating motion vectors, is generally considered the most expensive part of the encoding process. Similarly, motion estimation is likely to be the most expensive part of a transcoding process in which a video sequence is decoded, and then re-encoded with new parameters. With the advent of high-definition television ("HDTV"), this is a significant concern since television studios will be required to transcode MPEG-2 bitstreams of picture data from one format to another. For example, television studios accommodating the HDTV standard must be able to transcode an MPEG-2 bitstream from one group of pictures ("GOP") structure to another, from one picture size to another, and/or from one bit rate to another. There may be instances where these studios will need to transcode MPEG-2 pictures from field pictures to frame pictures, or from frame pictures to field pictures. There may also be instances where the studios will need to transcode MPEG-2 pictures from an interlaced sequence to a progressive sequence, or from a progressive sequence to an interlaced sequence. Transcoding may also be used to perform a "cropping" function wherein an input represents a main image, and the transcoded output represents a sub-image within the main image.

SUMMARY OF THE INVENTION

Therefore, it is an object of the present invention to provide a technique for transcoding a video sequence that minimizes or eliminates motion estimation.

It is another object to provide a technique for transcoding a video sequence in which motion information from an input bitstream is utilized to generate motion vectors for a transcoded output bitstream.

It is still another object to provide a motion estimation technique for the transcoding process that supports all of the MPEG-2 prediction modes.

These and other objects can be achieved in accordance with the principles of the present invention with a transcoding method performed by receiving a first bitstream of compressed image data having identifiable coding parameters. These parameters may relate to the GOP structure of pictures represented in the first bitstream, the size of the pictures represented in the first bitstream, whether the pictures represented in the first bitstream are field or frame pictures, and/or whether the pictures represented in the first bitstream define a progressive or interlaced sequence. First motion information is obtained from the first bitstream, and is used to extrapolate second motion information for a second bitstream of compressed image data. The second bitstream, which has one or more parameters different from the parameters of the first bitstream, is provided as a transcoded output.

BRIEF DESCRIPTION OF THE DRAWINGS

A more complete appreciation of this invention, and many of the attendant advantages thereof, will be readily apparent as the same becomes better understood by reference to the following detailed description when considered in conjunction with the accompanying drawings, wherein:

FIG. 1 illustrates the basic principles of motion vector extrapolation;

FIG. 2 illustrates a transcoding system constructed according to the principles of the present invention;

FIG. 3 illustrates the process of transcoding pictures from one GOP structure to another according to the principles of the present invention; and

FIG. 4 illustrates the general case of motion vector extrapolation, and can be used to better understand motion vector extrapolation as applied to transcoding.

DETAILED DESCRIPTION OF THE INVENTION

Turning now to the drawings and referring to FIG. 1, the basic principles of motion vector extrapolation are illustrated. In general, motion vector extrapolation is a technique for estimating motion for an image in a video sequence when motion is known elsewhere in the sequence. In FIG. 1, the bold-typed square in picture B represents an individual block of pixels (a.k.a., pels) within a video picture. Each block within the picture is assumed to have a fixed position. In MPEG-2, these blocks are referred to as "macroblocks", and each one consists of a 16x16 block of pixels. The solid arrow in FIG. 1 represents the known motion of the bold-typed block as the block moves from picture A to picture B. Actually, it is the imagery within the block that moves, not the block itself, which is in a fixed position. This known motion, defined by a motion vector, is used to extrapolate the block's motion as it moves from picture B to picture C. The dashed arrow in FIG. 1 represents the extrapolated motion vector. The arrows in FIG. 1 actually point in a direction opposite to the motion since FIG. 1 (as well as the other attached drawings)

use the more common coding convention in which the vector associated with a block points to the "source" of that block in the reference picture. In practice, a given motion vector is represented as a set of x and y coordinates which respectively indicate the horizontal and vertical motion of a block between two pictures. Extrapolation generally assumes that motion is uniform in time and "smooth" in space. Accordingly, the block in picture B is expected to move to the position shown in picture C, and all blocks it overlaps are expected to have approximately the same motion.

Referring now to FIG. 2, a transcoding system constructed according to the principles of the present invention is shown. In FIG. 2, the transcoding system 200 receives an input bitstream 201 of MPEG-2 compressed image data. This input bitstream 201, which exhibits a particular bit rate, represents pictures of a given size and GOP structure. The picture data of the input bitstream 201 may consist of field or frame pictures, and may represent a progressive or interlaced sequence. An MPEG-2 decoder 202, which is known in the art, receives the input bitstream 201 and decodes the same to generate an output of decompressed, digital image data 203. Decoder 202 also provides the motion information (*i.e.*, motion vectors) from the input bitstream 201 as a separate output 204. The digital image data 203 is input to an intermediate processing unit 205, which operates in the pixel domain and is capable of resizing the pictures represented by the digital image data 203 in a known manner. In addition to picture resizing, the intermediate processing unit 205 can perform the known cropping function by isolating a sub-image within the main image represented by the digital image data 203, and

providing output of data representative of the sub-image. The intermediate processing unit 205 is also capable of performing known interlacing and de-interlacing operations, which are necessary when transcoding a progressive sequence of pictures into an interlaced sequence of pictures, and when transcoding an interlaced sequence of pictures into a progressive sequence of pictures. The intermediate processing unit 205 performs its operations in response to an external input 206 that indicates which (if any) of the processing operations is to be performed, and further indicates the parameters to be used in such processing. Since the operations of the intermediate processing unit 205 are optional, if picture resizing, cropping, interlacing and de-interlacing are not performed for a given transcoding application, the intermediate processing unit 205 simply operates as a pass-through unit, and provides output of digital image data 207 without any pixel domain processing. In this case, digital image data 207 and digital image data 203 are obviously the same. Alternatively, if one or more of these processing operations is to be performed for a given transcoding application, the intermediate processing unit 205 performs the pixel domain processing operation(s) and provides output of the resulting processed, digital image data 207. The digital image data 207, whether or not processed by intermediate processing unit 205, is input to an MPEG-2 encoder 208 which re-encodes the received image data 207 with new parameters indicated by an external input 209. For example, the encoder 208 may encode the digital image data 207 to represent pictures of a GOP structure different from the GOP structure of the pictures represented by the input bitstream 201, and/or may adjust the bit rate. The encoder 208 may also encode the digital image data 207 as field or frame pictures. To perform the encoding

process, encoder 208 uses the motion information in output 204 from decoder 202 to extrapolate motion vectors for the pictures being re-encoded. Thereafter, the re-encoded pictures are provided from the transcoding system 200 as an output bitstream 210 of MPEG-2 compressed, digital image data.

As previously indicated, this output bitstream 210 may represent a picture size and GOP structure different from the size and GOP structure of the input bitstream 201, and the bit rate of the output bitstream 210 may be controlled to be different from the bit rate of the input bitstream 201. Moreover, the output bitstream 210 may represent field pictures while the input bitstream 201 represents frame pictures, or the output bitstream 210 may represent frame pictures while the input bitstream 201 represents field pictures. Similarly, the output bitstream 210 may represent an interlaced sequence while the input bitstream 201 represents a progressive sequence, or the output bitstream 210 may represent a progressive sequence while the input bitstream 201 represents an interlaced sequence. Also, the output bitstream 210 may represent a "cropped" version of the input bitstream 201. In essence, the transcoding system 200 receives the input bitstream 201 having certain parameters, and transcodes the input bitstream 201 to generate the output bitstream 210 having one or more parameters different from the parameters of the input bitstream 201. This transcoding system 200 is unique in that it uses motion information from the input bitstream 201 to extrapolate motion information for the output bitstream 210.

Turning now to FIG. 3, the process of transcoding pictures from one GOP structure to another according to the principles of the present invention is illustrated. In FIG. 3, an input bitstream represents pictures having a GOP structure of: an intra-coded picture ("I-picture"), a bidirectionally-predictive coded picture ("B-picture"), a B-picture, and a predictive-coded picture ("P-picture"). That is, the input bitstream represents an "IBBP" GOP structure. As indicated in FIG. 3, the two successive B-pictures of the input bitstream are labeled as B1 and B2 for clarity. This input bitstream is analogous to the input bitstream 201 in FIG. 2. During the transcoding process, the input bitstream having an "IBBP" GOP structure is converted to an output bitstream having an "IPPP" GOP structure. As indicated in FIG. 3, the three successive P-pictures of the output bitstream are labeled as P1, P2 and P3 for clarity. This output bitstream is analogous to the output bitstream 210 in FIG. 2. To minimize clutter, the macroblocks that make up the individual pictures are not shown in FIG. 3. To perform the transcoding process, the I-P and P-B2 motion vectors from the input bitstream are utilized to extrapolate motion vectors for the output bitstream. For example, as FIG. 3 indicates, the I-P motion vector from the input bitstream can be used to estimate a P1-P2 motion vector for the output bitstream. Similarly, the P-B2 motion vector from the input bitstream can be used to estimate a P2-P3 motion vector for the output bitstream.

A more detailed explanation of the motion vector extrapolation process of the present invention will now be provided with reference to FIG. 4, which shows four pictures in a video sequence. Conceptually, these four pictures will exist in the output

bitstream as well as the input bitstream, and this will usually be the case in fact. In any case, in FIG. 4, picture T is the "target" picture whose motion relative to picture R is to be determined. Both of these pictures will exist in the output bitstream. Pictures KR and KT are pictures whose relative motion is known from the input bitstream. According to the principles of the present invention, the known motion between pictures KR and KT can be used to extrapolate motion vectors for the output bitstream. Pictures KR and KT are referred to as a "base pair", and pictures R and T are referred to as a "current pair". Several base pairs may be useful to estimate the motion for a current pair. In general, any base pair temporally close to the current pair is potentially useful for this purpose.

During the transcoding process, there may be instances where some blocks in picture T do not receive motion extrapolation information. That is, there may be instances where some blocks in picture T are not overlapped by the motion of any blocks in nearby base pairs. Such blocks in picture T can simply be intra-coded, or their motion can be calculated in a conventional manner. Alternatively, the motion of such blocks can be inferred or interpolated from nearby blocks. In most cases, however, several blocks in base pairs will overlap most of the blocks in picture T, and the task is then to select the best motion vector from among several candidate motion vectors. Each overlapping block from a base pair provides a single candidate motion vector. When selecting from among multiple candidate motion vectors, it is useful to attach a weight to each motion vector, and select the motion vector having the greatest weight. Using the pictures of FIG. 4 as an example, the weight for each candidate motion vector (MV) is calculated as follows:

$$\text{Total Weight (MV)} = (\text{Overlap Weight}) * (\text{Picture Weight}) * (\text{Rounding Weight}) \dots (1)$$

where,

$$\text{Overlap Weight} = (\text{the number of pixels in a block of picture KT that overlap} \dots \text{pixels in the corresponding block of picture T}) \dots (2)$$

$$\text{Picture Weight} = 1/[1+\text{abs}(\text{Temporal Weight})] \dots (3)$$

where,

$$\text{Temporal Weight} = (t_{KT}-t_{KR}) * [(t_{KR}+t_{KT})-(t_T+t_R)] \dots (4)$$

and,

$$\text{Rounding Weight} = [1-(\text{horizontal rounding error})] * [1-(\text{vertical rounding error})] \dots (5)$$

In the foregoing equations (1) through (5), t_{KR} , t_{KT} , t_R and t_T represent the display times of pictures KR, KT, R and T, respectively. The terms * and abs represent the multiplication operator and the absolute value designation, respectively. The horizontal and vertical rounding errors occur when the extrapolated vector is rounded to the nearest half pixel, and accordingly, each takes on a value from zero (0) to one-half (1/2). Although equations (1) through (5) are shown as applied to the pictures of FIG. 4, it should be intuitive that these general equations can be applied to other picture configurations.

In addition to the technique of simply selecting the candidate motion vector having the greatest weight, the present invention contemplates other ways of obtaining the best motion vector(s) for a given block. According to an alternative embodiment, one can calculate the weighted average where the weights are computed using the foregoing

equations (1) through (5). Using these weights, the best motion vector can then be calculated on a component basis through multiplying the weights by the x or y components of the corresponding motion vectors to generate weighted components, summing the weighted components, and dividing the sum of the weighted components by a sum of the weights.

The foregoing weighted average technique can also be performed on a "cluster" basis. That is, if an x-y plotting of the candidate motion vectors shows more than one cluster (i.e., close grouping) of motion vectors, a best motion vector can be calculated for each cluster. This is the centroid, or "center of mass" of the cluster. A final best motion vector can then be selected from the best motion vectors of the individual clusters.

It should be noted that a best motion vector can also be selected from among a plurality of candidate motion vectors without using weights at all. For example, mean-square error (MSE) or mean-absolute-difference (MAD) calculations representing pixel differences between blocks can be performed to identify the best motion vector. The use of MSE and MAD calculations is, of course, well known to those skilled in the art.

The following algorithm listed in Table 1 shows the steps for finding the best motion vector for each possible prediction mode of a target picture (e.g., picture T in FIG. 4). These steps, which are represented by pseudo-code, can be programmed in any programming language known to those skilled in the art.

 Table 1: Pseudo-Code For Finding The Best Motion Vector

- for each prediction mode to be considered for target picture (current pair)
 - for each field or frame of the applicable mode (top field, bottom field, frame)
 - initialize a table indexed by blocks in target picture
 - for each base pair considered useful for a current pair (e.g., a base pair is generally considered "useful" if it is temporally nearby within a preset range)
 - for each motion vector of each inter-coded block in the target of the base pair
 - determine where the vector moves the block to in the target picture
 - calculate the vector's weight, or evaluate its match (e.g., MSE, MAD)
 - store info in table entries of relevant target picture blocks
 - for each block in target picture
 - determine best motion vector for field or frame
- for each block in target picture
 - determine best prediction mode and corresponding motion vector(s)

Note that multiple motion vectors are associated with a given block in some prediction modes, such as field prediction for frame pictures and dual prime for P-pictures. Also, for B-pictures, the encoder must decide for each block whether to use forward prediction, backward prediction or both. In some cases no prediction mode

produces acceptable results. This occurs when there are no candidate motion vectors, or when the best motion vector determined in accordance within one of the foregoing techniques is not good enough for the particular application. Additionally, there are instances where the encoder can simply use the motion vectors (unchanged) from the input sequence for the output sequence.

Assume in FIG. 4 a situation using frame prediction for frame pictures, with no resizing. Given that a block in picture KT has an associated vector, v_K , then one or more blocks in picture T will have an associated vector, v , defined as follows:

$$v = T_v * v_K \dots\dots\dots (6)$$

where,

$$T_v = (t_T - t_R) / (t_{KT} - t_{KR}) \dots\dots\dots (7)$$

Here, T_v is the temporal scaling factor for vectors, and t_{KR} , t_{KT} , t_R and t_T represent the display times of pictures KR, KT, R and T, respectively. There is an ambiguity in the display times for frame pictures whose fields are displayed at different times. In such cases, the times should be the averages of the respective fields. To track the motion of a block from picture KT to picture T, consider a point at the upper left hand corner of the particular block. In the context of FIG. 4, this point is represented as point q_K in picture KT, and as point q in picture T. The motion of the point q is given by:

$$q = q_K - (T_b * v_K) \dots\dots\dots (8)$$

where,

$$T_b = (t_T - t_{KT}) / (t_{KT} - t_{KR}) \dots\dots\dots (9)$$

In a transcoding process involving resized pictures, pictures KR and KT will not have the same size as pictures R and T. Resizing is accommodated as follows. Let Rx and Ry represent the horizontal and vertical enlargements of the pictures, respectively. That is:

$$R_x = (\text{width of picture T})/(\text{width of picture KT}) \dots\dots\dots (10)$$

and,

$$R_y = (\text{height of picture T})/(\text{height of picture KT}) \dots\dots\dots (11)$$

The resizing of pictures requires equations (6) and (8) to be modified as follows:

$$v = R(Tv * v_K) \dots\dots\dots (12)$$

and,

$$q = R(q_K - (Tb * v_K)) \dots\dots\dots (13)$$

The foregoing equations introduce the notation:

$$R(V) = (V_x * R_x, V_y * R_y) \dots\dots\dots (14)$$

where V is a two-component point (pixel) or vector. Note that R refers to different pictures sizes, not to different block sizes.

Resizing applies not only to resized images, but also to relating field and frame pictures, and to relating field and frame prediction. To accommodate field prediction (of both field and frame pictures), note that the bottom fields are 0.5 pixels lower than their coordinates would indicate. The displacement vector, dX, of a field or frame picture is defined as follows:

$$dX = (0,0) \text{ if } X \text{ is a top field or a frame} \dots\dots\dots (15)$$

and,

$$dX = (0,0.5) \text{ if } X \text{ is a bottom field} \dots\dots\dots(16)$$

Applying this to equations (12) and (13), the new equations for v and q become:

$$v = R(Tv*(vK+dKR-dKT))+dT-dR \dots\dots\dots(17)$$

and,

$$q = R(qK+dKT-[Tb*(vK+dKR-dKT)])-dT \dots\dots\dots(18)$$

Therefore, the final equations used in the practice of the present invention are equations (7), (9) and (14) through (18).

Finally, suppose the intermediate processing unit 205 of FIG. 2 crops the images. The above equations can be applied as follows. We consider pictures R and T to be the same size as the original, with the output images imbedded in them at the appropriate location. This location defines the region of interest. Only blocks that overlap the region of interest are used in extrapolation. Then, FIGS. 1 through 4, as well as the equations derived from them, apply to this case.

The foregoing disclosure may suggest homogeneity; that is, if picture T is a frame picture using field-based prediction, then picture KT is also a frame picture using field-based prediction. However, the previously discussed formulas are more general, as the following examples illustrate.

Suppose picture T uses frame-based prediction, and picture KT uses field-based prediction. In the present invention, it does not matter whether picture KT is a field or frame picture. In either case, picture KT's motion vectors are field based, so that $R_y = 2$ and $R_x = 1$. If picture KT is a frame picture, its blocks are 16x8; however, that is transparent to the foregoing formulas. If a block uses 16x8 prediction, then its two halves should be treated as two separate blocks.

Suppose next that both picture T and picture KT use field-based prediction. Here, since vectors in both pictures refer to images of the same size, resizing is not required and $R_x = R_y = 1$. That is, in the practice of the present invention it is irrelevant to the equations whether there is a mix of field and frame pictures, or of field and frame prediction.

As implied by the foregoing disclosure, if a block in picture KT uses dual-prime, then its two vectors are applied separately, perhaps with different weights. Also, evaluating picture T for dual-prime is much like evaluating the same for bi-directional interpolation. That is, each vector is evaluated separately, and then the combination is evaluated. Accordingly, the present invention supports all of the MPEG-2 prediction modes: frame, field and dual prime prediction for frame pictures; and field, dual prime and 16x8 prediction for field pictures.

It is noted that the equations disclosed herein are generally directed towards situations of linear motion, rather than non-linear motion. Non-linear motion can result, for example, from camera jitter or uniform acceleration of objects in pictures. Such instances of non-linear motion can be handled as follows. For non-linear motion associated with camera jitter, it may be desirable to factor out the global motion before comparing block motion. The known MSE and MAD calculations can be used for this purpose. For non-linear motion associated with the uniform acceleration of objects, it may be desirable to extrapolate this motion using three pictures, rather than two.

While there have been illustrated and described what are considered to be preferred embodiments of the present invention, it will be understood by those skilled in the art that various changes and modifications may be made, and equivalents may be substituted for elements thereof without departing from the true scope of the present invention. In addition, many modifications may be made without departing from the central scope thereof. Therefore, it is intended that the present invention not be limited to the particular embodiment disclosed as the best mode contemplated for carrying out the invention, but that the present invention includes all embodiments falling within the scope of the appended claims.

What is claimed is:

1. A transcoding method, comprising the steps of:
receiving a first bitstream of compressed image data having a first GOP structure;
obtaining first motion information from the first bitstream;
utilizing the first motion information to extrapolate second motion information for
a second bitstream of compressed image data having a second GOP structure different
from the first GOP structure; and
providing output of the second bitstream.
2. The method as claimed in claim 1, further comprising a step of controlling
a bit rate of the second bitstream so that a bit rate of the first bitstream is different from
the bit rate of the second bitstream.
3. The method as claimed in claim 1, further comprising a step of adjusting a
size of pictures represented by the first bitstream so that pictures represented by the
second bitstream exhibit a size different from the size of the pictures represented by the
first bitstream.
4. The method as claimed in claim 2, further comprising a step of adjusting a
size of pictures represented by the first bitstream so that pictures represented by the

second bitstream exhibit a size different from the size of the pictures represented by the first bitstream.

5. The method as claimed in claim 4, further comprising a step of encoding the pictures represented by the second bitstream as field pictures when the pictures represented by the first bitstream are encoded as frame pictures.

6. The method as claimed in claim 4, further comprising a step of encoding the pictures represented by the second bitstream as frame pictures when the pictures represented by the first bitstream are encoded as field pictures.

7. The method as claimed in claim 4, further comprising a step of interlacing the pictures represented by the first bitstream when the pictures represented by the first bitstream are received as a progressive sequence so that the pictures represented by the second bitstream are output as an interlaced sequence.

8. The method as claimed in claim 4, further comprising a step of de-interlacing the pictures represented by the first bitstream when the pictures represented by the first bitstream are received as an interlaced sequence so that the pictures represented by the second bitstream are output as a progressive sequence.

9. The method as claimed in claim 1, further comprising a step of encoding pictures represented by the second bitstream as field pictures when pictures represented by the first bitstream are encoded as frame pictures.

10. The method as claimed in claim 1, further comprising a step of encoding pictures represented by the second bitstream as frame pictures when pictures represented by the first bitstream are encoded as field pictures.

11. The method as claimed in claim 1, further comprising a step of interlacing pictures represented by the first bitstream when the pictures represented by the first bitstream are received as a progressive sequence so that pictures represented by the second bitstream are output as an interlaced sequence.

12. The method as claimed in claim 1, further comprising a step of de-interlacing pictures represented by the first bitstream when the pictures represented by the first bitstream are received as an interlaced sequence so that pictures represented by the second bitstream are output as a progressive sequence.

13. A transcoding method, comprising the steps of:
receiving a first bitstream of compressed image data representing pictures of a first size;

obtaining first motion information from the first bitstream;

utilizing the first motion information to extrapolate second motion information for a second bitstream of compressed image data representing pictures of a second size different from the first size; and

providing output of the second bitstream.

14. A transcoding method, comprising the steps of:

receiving a first bitstream of compressed image data representing pictures defining an interlaced sequence;

obtaining first motion information from the first bitstream;

utilizing the first motion information to extrapolate second motion information for a second bitstream of compressed image data representing pictures defining a progressive sequence; and

providing output of the second bitstream.

15. A transcoding method, comprising the steps of:

receiving a first bitstream of compressed image data representing pictures defining a progressive sequence;

obtaining first motion information from the first bitstream;

utilizing the first motion information to extrapolate second motion information for a second bitstream of compressed image data representing pictures defining an interlaced sequence; and

providing output of the second bitstream.

16. A transcoding method, comprising the steps of:
receiving a first bitstream of compressed image data representing frame pictures;
obtaining first motion information from the first bitstream;
utilizing the first motion information to extrapolate second motion information for
a second bitstream of compressed image data representing field pictures; and
providing output of the second bitstream.

17. A transcoding method, comprising the steps of:
receiving a first bitstream of compressed image data representing field pictures;
obtaining first motion information from the first bitstream;
utilizing the first motion information to extrapolate second motion information for
a second bitstream of compressed image data representing frame pictures; and
providing output of the second bitstream.

18. A transcoding method, comprising the steps of:
receiving a first bitstream of compressed image data representing a main image;
obtaining first motion information from the first bitstream;
utilizing the first motion information to extrapolate second motion information for
a second bitstream of compressed image data representing a portion of the main image;
and
providing output of the second bitstream.

19. A transcoding method, comprising the steps of:

receiving a first bitstream of compressed image data having a plurality of coding parameters including a GOP structure, a picture size, a bit rate, a frame or field picture format, and a progressive or interlaced sequence;

obtaining first motion information from the first bitstream;

utilizing the first motion information to extrapolate second motion information for a second bitstream of compressed image data having the plurality of coding parameters such that one or more of the coding parameters of the second bitstream are different from the coding parameters of the first bitstream; and

providing output of the second bitstream.

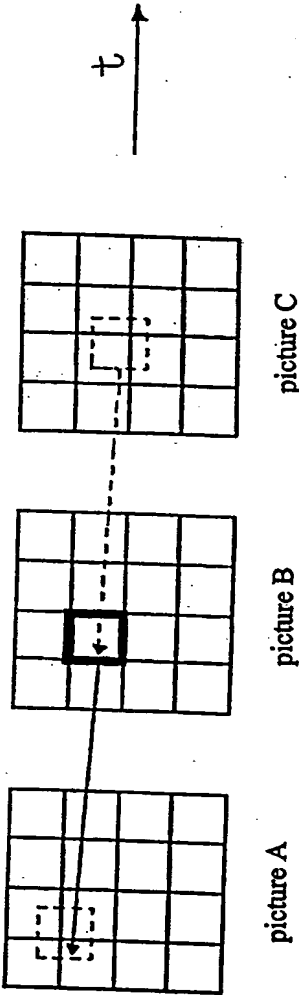


FIG. 1

200

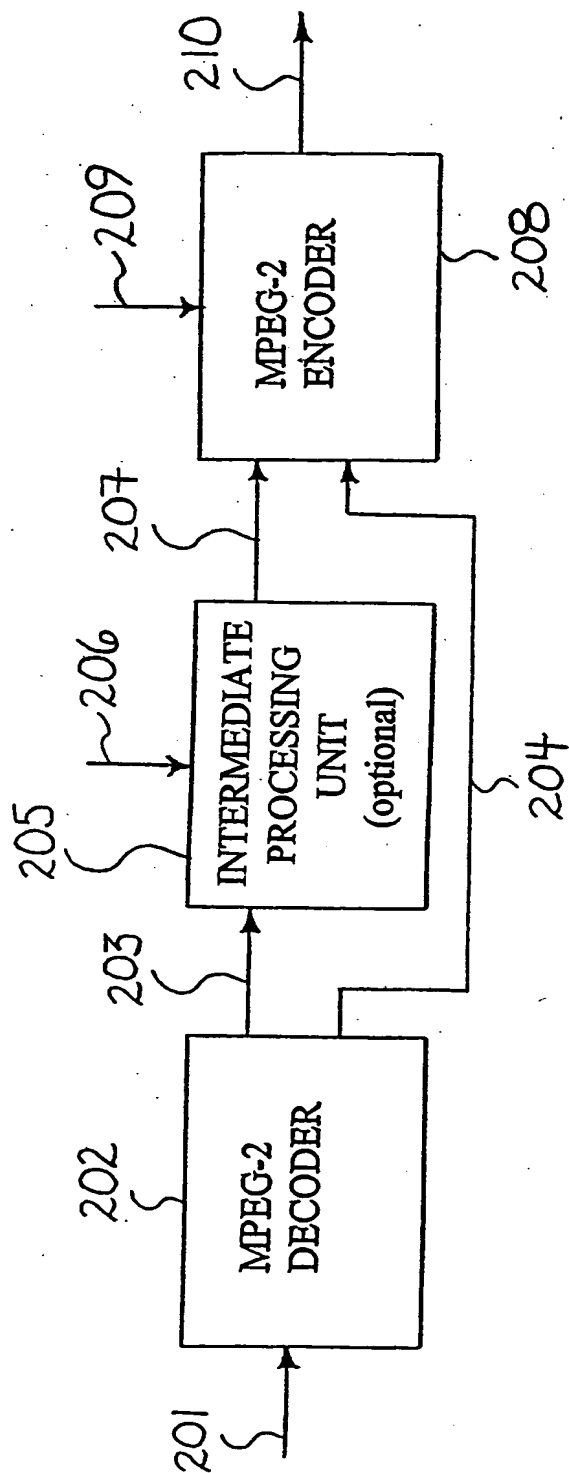


FIG. 2

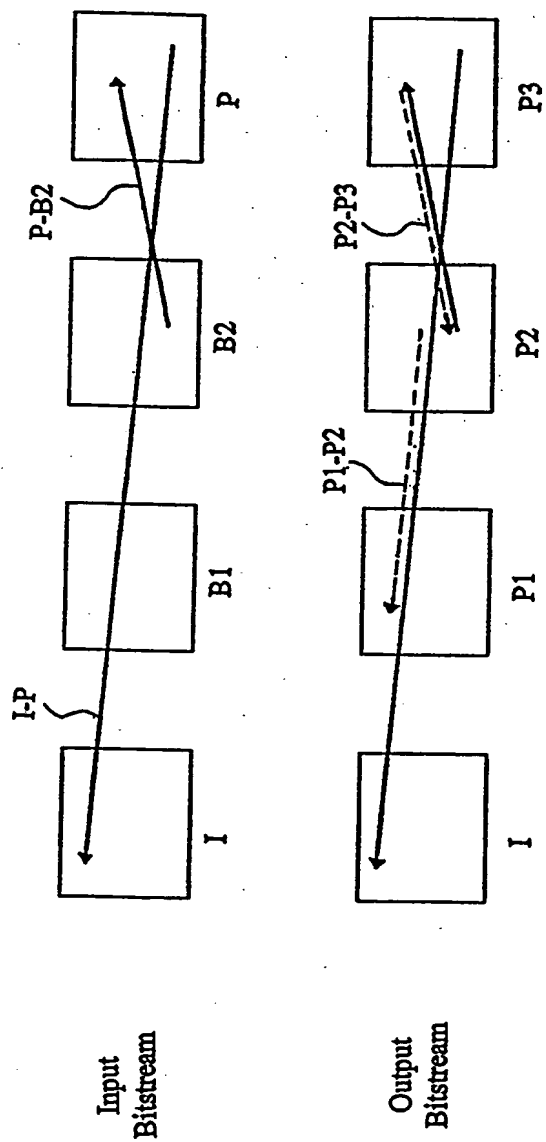


FIG. 3

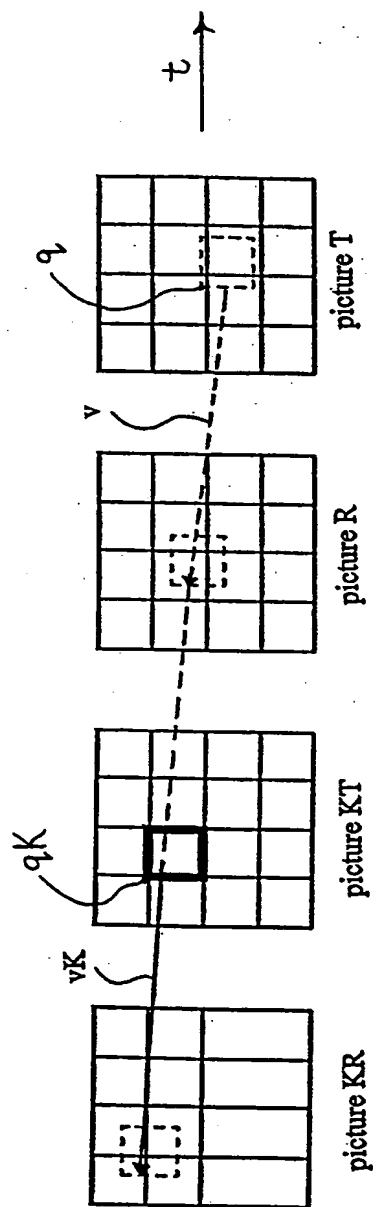


FIG. 4

INTERNATIONAL SEARCH REPORT

International application No

PCT/US 99/02318

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 H04N7/26 H04N7/36

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category * | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|------------|---|--|
| X Y | <p>EP 0 637 893 A (NEDERLAND PTT) 8 February 1995 see abstract</p> <p>see column 7, line 4 - line 26 see column 3, line 37 - line 40 see column 3, line 51 - line 57 see column 9, line 25 - line 32 see figure 1</p> <p style="text-align: center;">--- -/--</p> | <p>1-4, 13, 19 5-12, 14-17</p> |

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "A" document member of the same patent family

Date of the actual completion of the international search

10 May 1999

Date of mailing of the international search report

18/05/1999

Name and mailing address of the ISA
European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Berbain, F

INTERNATIONAL SEARCH REPORT

Internatio Application No
PCT/US 99/02318

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

| Category * | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|------------|--|-------------------------|
| Y | STOETTELDER U ET AL: "BRIDGING STANDARDS FOR DIGITAL TELEVISION BROADCAST" IEEE TRANSACTIONS ON CONSUMER ELECTRONICS, vol. 41, no. 3, 1 August 1995, pages 626-633, XP000539516 | 5-12, 14-17 |
| A | see abstract see paragraph 3 see figures 1,3,4 | 1-4,13, 18,19 |
| X | SHEN B ET AL: "ADAPTIVE MOTION VECTOR - RESAMPLING FOR COMPRESSED VIDEO DOWN -SCALING" PROCEEDINGS INTERNATIONAL CONFERENCE ON IMAGE PROCESSING, 26 October 1997, pages 771-774, XP000792878 see abstract see paragraph 2 | 13,18,19 |
| A | see figure 3 | 1-12,14, 17 |
| P,X | WO 98 19460 A (CHRISTOPOULOS CHARILAOS ;ERICSSON TELEFON AB L M (SE); BJOERK NIKL) 7 May 1998 see abstract see figure 3 see page 13, line 15 - line 25 see page 19, line 1 - line 20 | 1-4,13, 19 |
| A | WO 96 23388 A (POLOMSKI MARK D ;VIDEO SERVER INC (US)) 1 August 1996 see page 35, line 1 - page 36, line 20 see page 37, line 17 - line 30 see figure 25 | 1-12, 14-18 13,19 |
| X | | |
| A | EP 0 696 873 A (AT & T CORP) 14 February 1996 see abstract see column 7, line 4 - line 22 see figure 3 | 1-19 |

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 99/02318

| Patent document cited in search report | Publication date | Patent family member(s) | Publication date |
|---|---------------------|----------------------------|---------------------|
| EP 0637893 A | 08-02-1995 | NL 9301358 A | 01-03-1995 |
| | | AT 174178 T | 15-12-1998 |
| | | AU 672507 B | 03-10-1996 |
| | | AU 6880594 A | 16-02-1995 |
| | | CA 2128793 A,C | 05-02-1995 |
| | | CN 1111878 A | 15-11-1995 |
| | | DE 69414931 D | 14-01-1999 |
| | | FI 943622 A | 05-02-1995 |
| | | JP 2860880 B | 24-02-1999 |
| | | JP 7095090 A | 07-04-1995 |
| | | KR 127094 B | 29-12-1997 |
| | | NO 942866 A | 06-02-1995 |
| | | US 5544266 A | 06-08-1996 |
| WO 9819460 A | 07-05-1998 | AU 4798097 A | 22-05-1998 |
| | | SE 9603909 A | 26-04-1998 |
| WO 9623388 A | 01-08-1996 | US 5600646 A | 04-02-1997 |
| | | AU 4698796 A | 14-08-1996 |
| | | DE 19681223 T | 08-01-1998 |
| | | GB 2312807 A | 05-11-1997 |
| EP 0696873 A | 14-02-1996 | US 5512953 A | 30-04-1996 |
| | | CA 2152330 A | 10-02-1996 |
| | | JP 8070440 A | 12-03-1996 |